Comparative Analysis of Uncertainty Characterization Methods in Urban Building Energy Models in Hot-Arid Regions

Dongxue Zhan¹, Nurettin Sezer², Ibrahim Galal Hassan², Liangzhu (Leon) Wang¹
¹ Centre for Zero Energy Building Studies, Department of Building, Civil and Environmental Engineering, Concordia University, Montreal, Quebec, Canada
² Mechanical Engineering Program, Texas A&M University at Qatar, Engineering Building, Education City, Doha, Qatar

Abstract
The development of reliable building energy models at the urban scale is crucial for analyzing and optimizing the energy efficiency of cities. The bottom-up physics-based approach has been widely employed in Urban Building Energy Models (UBEMs). However, the uncertainty of input parameters can impact the reliability of UBEM simulation outputs, and very limited studies considered the uncertainty when developing archetype models for UBEMs. While UBEMs typically rely on a traditional deterministic approach, incorporating probabilistic methods can significantly enhance simulation accuracy by accounting for uncertain variables. Probabilistic methods involve characterizing key uncertainties in input data using Probability Distribution Functions (PDFs). Yet, the effect of using different PDF types on UBEM results is not adequately understood, and the literature often assumes uniform distribution. In this study, UBEM is characterized based on three methods. The deterministic approach is used to serve as a baseline, and two different PDF types are used to examine how PDFs impact simulation results when uncertain parameters are present in UBEMs. Latin Hypercube Sampling (LHS) is employed to propagate uncertainty in input parameters in UBEM. The study is conducted on a case study area of the Marina district of Lusail City, Qatar, characterized by a hot and arid climate.

Highlights
- A comparative analysis of deterministic and probabilistic approaches to urban building energy models is conducted.
- A total of 1600 UBEM simulations were performed to generate probabilistic results.
- Ten uncertain parameters were included in the probabilistic models.
- The impact of two different probability distribution functions on UBEM was compared.

Introduction
The building sector is a significant contributor to global energy consumption and greenhouse gas emissions, with buildings consuming 36% of global energy and constituting 39% of energy-related carbon dioxide emissions, according to the International Energy Agency (IEA) (IEA, 2019). Improving the energy efficiency of buildings is crucial to reducing global energy demand and mitigating climate change.

Building Energy Modeling (BEM) is a powerful tool to predict and optimize buildings’ energy performance. UBEM allows for identifying energy-saving potentials and improving energy efficiency in large-building stocks in cities, leading to reduced energy costs and greenhouse gas emissions and supporting sustainable development. Considerable research has been devoted to the development of bottom-up, physics-based UBEM tools, which enable the evaluation of district-wide energy demand and supply strategies. To reduce the number of simulation inputs required in UBEM, groups of buildings are often classified into representative “archetypes”. This approach simplifies the actual diversity of occupant behaviors and construction variations, reducing the variety of predicted building energy demands. However, it is important to note that this simplification is necessary for efficient and effective UBEM simulations. The accuracy of these simulations can be improved by using more detailed inputs and accounting for the diversity of building characteristics and occupant behaviors. The UBEM inputs cover building geometry, envelope material properties, schedules of HVAC systems and electric equipment, as well as local climate conditions.

In BEM, building archetypes are used to simplify the assignment of non-geometric simulation variables to individual building models. To this end, buildings are grouped based on one or more categories or indicators that are correlated with the energy demand of the building and are available for all buildings. The most common indicators used to classify buildings into archetypes are building type, floor area, and age of construction. Once all the buildings have been classified, each resulting archetype has to be characterized for all relevant energy simulation parameters. The most common approach is to characterize archetype parameters in a deterministic way, assigning a single value to each parameter that is used for each building. However, the lack of benchmarking ordinances and laws in some urban areas, including some of those in hot-arid regions, makes it difficult to accurately describe the input parameters required for UBEMs. This introduces uncertainty and unreliability, which can hinder the effective use of UBEMs. The deterministic approach for archetype model may fail to accurately represent the buildings. Furthermore, even if the archetype building correctly represents the mean of all buildings in its group, individual buildings will necessarily perform differently. To account for uncertainties regarding simulation inputs, these inputs can
be characterized in a probabilistic way by defining them as distributions. The lack of sufficient information for each input parameter introduces uncertainty and unreliability, which is one of the major barriers to the effective use of UBE Ms. In this regard, Uncertainty Analysis (UA) has been widely used to characterize input uncertainties and evaluate the robustness and reliability of models. UA becomes more crucial in UBE M since the knowledge of single building becomes imprecise on a large scale when the traditional building simulation is scaled up to the urban scenarios. UA has been reported to improve the accuracy of UBE Ms for energy demand and peak power estimation (Prataviera et al., 2022).

In the probabilistic approach, input parameters are typically associated with Probability Distribution Functions (PDFs), and various sampling techniques are used to generate large input datasets for UBE Ms. A variety of sampling techniques have been used in building energy analysis, such as Standard Monte Carlo sampling (Prataviera et al., 2022, Y. Chen et al., 2020) and Latin Hypercube Sampling (LHS) (X. Chen et al., 2017; Parsy et al., 2012). To obtain combinations of input variables based on PDFs, LHS is widely used in building performance analysis due to its efficient stratification properties. Due to its efficient stratification properties, the LHS method is an efficient alternative to commonly used methods, such as Standard Monte Carlo, to account for uncertainty propagation in building performance analysis (Tian, 2013). However, most researchers have either relied on an assumption of uniform distribution or used certain PDFs with limited reference. For instance, Kong et al. (Kong et al., 2023) assigned uniform probability distribution to all 12 weather-unrelated input variables in building energy simulation for an underground metro station. Cerezo et al. (Cerezo et al., 2017) also assigned uniform probability distribution to occupancy, lighting power density, plug multiplier, and cooling set point in uncertainty analysis. Further, Chen et al. (Y. Chen et al., 2020) assigned seventeen input variables selected with four value options in the uncertainty UBEM model.

Some literature assigned certain PDFs to UBE M input parameters. Prataviera et al. (Prataviera et al., 2022) assigned generalized extreme value, triangular, and uniform PDFs to operational, geometrical, and physical uncertain parameters according to the Italian national survey and scientific literature. In addition, Alghamdi et al. (Alghamdi et al., 2022) assigned normal probability distribution to all the input parameters, including window-to-wall ratio, cooling/heating set point temperature, building orientation, occupancy density, ventilation rate, thermal mass, roof window open ratio, and infiltration. Detailed parameter probability distribution was assigned to each parameters based on literature, theoretical considerations, and educated guesses (Domínguez-Muñoz et al., 2010).

To date, there has been very limited investigation into the impact of selecting a distribution for each input parameter on the simulation results of UBE Ms. However, many scholars rely on assumed uniform distributions or use certain distribution PDFs with limited reference without investigating the impact of selecting a distribution for each input parameter on the simulation results of UBE Ms. Consequently, more research is needed to explore the effect of different probability distributions on UBE M results.

This study aims to quantify the impact of incorporating uncertainties into UBE M development using a probabilistic approach compared to the traditional deterministic approach. Additionally, the effects of two different types of PDFs on the results of the UBE M are compared. The case study is conducted in the Marina district of Lusail City, Qatar, to gain a better understanding of the influence of PDF selection on the propagation of uncertainties in UBE M.

**Methodology**

- **Methodology overview**

  The methodology is outlined in Figure 1. It begins with data collection from standards and literature to gather input parameters. Next, parameter characterization is performed using deterministic and PDF methods. Random samples of parameters are generated using the LHS method. The input datasets are then modified accordingly, followed by UBE M simulations absorbing modified datasets to comprehensively assess energy demand and associated uncertainties. Finally, future retrofit actions can be guided by the results of the uncertainty analysis.

  The geometry data is essential in UBE M simulations. The accuracy and completeness of geometry data directly impact the results of UBE M simulations, as it defines the shape and size of the building. The GIS information of the case study was obtained using a method proposed by Katal et al. (Katal et al., 2022) for the 3D city generation by integrating publicly available data sets (OpenStreetMap and Microsoft footprints) and a free program (Google Earth). These data sets provide 2D building footprints, whereas Google Earth provides digital surface models of terrains and buildings. The developed geometry of the Marina district of Lusail City was further compared and confirmed with Qatar local documents “Lusail City Marina District Design Guidelines” and “Lusail City-Marina District Construction Map,” as well as with information obtained from an industrial partner. To exclude the uncertainties in meteorological variables in UBE M simulations, the weather data measured by a local weather station in the Marina district was implemented as weather inputs. The local weather station was installed at coordinates 25.399952 E and 51.519568 N. The sensors of the weather station report hourly data on air temperature, solar radiation, wind direction, wind speed, relative humidity, and precipitation.
Case study area

The overestimation of building cooling loads adversely impacts the cost of equipment selection, system sizing and design, and operation of district cooling plants. Accurate prediction of district cooling loads remains challenging in regions with hot climates such as Qatar. The Marina district is an ongoing project to become the future downtown of Lusail City in Qatar, which is under colossal construction. The Marina district is an interesting case study area for testing different uncertainty characterization approaches on UBEM development. The Marina district (Figure 2) covers a total of 3.1 million m² of built-up area. The population of this district is 40,760, including 27,000 residents. The district comprises 148 buildings, including 92 residential buildings and 56 office buildings.

Archetype characterization methods

Building archetype is an effective simplification tool in UBEM for assigning non-geometric simulation variables to individual building models. In this study, the buildings are categorized by the building use (office and residential) and age of construction. The entire Marina district was constructed in the last ten years. Therefore, two building archetypes are used in this study according to the classification considering building type and age of construction. Then, each resulting archetype has to be characterized for all relevant uncertain energy simulation parameters. These include the factors influencing building energy demand, including envelope composition, thermal properties, internal loads, occupancy, etc. Concerning envelope composition, the input parameters include wall insulation thermal resistance, roof insulation thermal resistance, roof solar reflectance, window overall heat transfer coefficient, and window solar heat gain coefficient (SHGC). A total of ten parameters are considered, as listed in Table 2.

All the datasets for archetype characterization in the case study are collected from available resources, including the related standards and literature (Table 1). The data for input parameters are categorized into three: building envelope, internal loads, and equipment, which are listed in detail in Qatar Construction Specifications (QCS), which is intended for use with the General Condition of Contract (Qatar Construction Specification, 2014). QCS describes the minimum envelope performance requirements. On the other hand, ASHRAE Standard 90.1 2019 (ANSI/ASHRAE/IES Standard 90.1-2019 -- Energy Standard for Building Except Low-Rise Residential Buildings, 2019) serves as a benchmark for commercial building energy codes in the United States. Besides, it serves as a foundation for the development of codes and standards worldwide. ASHRAE Standard 90.1 2019 also describes the building envelope requirements classified by climate zone for non-residential and residential buildings. Moreover, Global Sustainability Assessment System (GSAS) is a regionally recognized green building rating system developed by Gulf Organization for Research & Development (GORD). Further, Lusail City GSAS 2 Star Guidelines (Lusail City GSAS 2 Star Rating Guidelines.Pdf, 2014) provides the minimum requirement for Lusail City projects.

Three different archetype definition approaches are implemented in the comparative evaluation of the case study with an increasing level of detail for input parameters in UBEM. The first method, “Method 1” is the typical deterministic approach based on the standards and literature. While Methods 2 and 3 introduced LHS sampling methods for input parameters with uniform and certain PDFs, respectively. Table 2 summarizes ten uncertain factors identified in the model, the proposed probability distributions, and a deterministic value for each factor. The deterministic approach utilized values that were equivalent to the median value within the input range of the probabilistic approaches.

Furthermore, the value range for the two probabilistic approaches was identical. These distributions have been proposed based on literature (Domínguez-Muñoz et al., 2010; Prataviera et al., 2022), and the values are based on the standards listed in Table 1. In the table, a distribution type $U[a, b]$ is a uniform distribution between $a$ and $b$, while a distribution type $T[a, c, b]$ is a triangular distribution with minimum value $a$, maximum $b$, and mode $c$.

The LHS method is used to simulate the random variation in input parameters of UBEM, since it provides good convergence of parameters space with relatively fewer samples. LHS is designed to reduce the computational burden by discretizing the parameter space into intervals of equal probability. By selecting a single sample for each interval, the number of required simulations is reduced while still providing a representative sampling of the parameter space. There are 800 cases generated for methods 2 and 3 for each, with ten parameters included, the detailed distribution of input parameters is listed in Table 2. Totally 1600 UBEM simulations are conducted to examine the impact of different PDFs on building hourly cooling energy consumption.

![Figure 1: Schematic illustration of the methodology](image-url)
Table 1: List of input parameters and data sources for prototype building model development

<table>
<thead>
<tr>
<th>Category</th>
<th>Parameter</th>
<th>Source of data</th>
</tr>
</thead>
</table>

Table 2: Summary of prototype input parameter values for deterministic models and probability models

<table>
<thead>
<tr>
<th>ID</th>
<th>Parameter (Unit)</th>
<th>Type</th>
<th>Deterministic value</th>
<th>Probability distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>WALU</td>
<td>Wall U value (W/m²K)</td>
<td>All</td>
<td>0.5</td>
<td>T[0.3, 0.5, 0.7]</td>
</tr>
<tr>
<td>ROFU</td>
<td>Roof U value (W/m²K)</td>
<td>All</td>
<td>0.3</td>
<td>T[0.19,0.3, 0.44]</td>
</tr>
<tr>
<td>WINU</td>
<td>Window U value (W/m²K)</td>
<td>All</td>
<td>2.3</td>
<td>T[1.8, 2.3, 2.8]</td>
</tr>
<tr>
<td>SHGC</td>
<td>Window SHGC</td>
<td>All</td>
<td>0.26</td>
<td>U[0.25, 0.275]</td>
</tr>
<tr>
<td>OCD</td>
<td>Occupancy density (Person/100m²)</td>
<td>COM</td>
<td>32.5</td>
<td>U[5, 60]</td>
</tr>
<tr>
<td>APD</td>
<td>Appliance power density (W/m²)</td>
<td>COM</td>
<td>12</td>
<td>T[2.5, 12, 21.53]</td>
</tr>
<tr>
<td>LPD</td>
<td>Lighting power density (W/m²)</td>
<td>COM</td>
<td>9</td>
<td>T[8.5, 9, 9.6]</td>
</tr>
<tr>
<td>VENT</td>
<td>Ventilation rate (cfm/ft²)</td>
<td></td>
<td>0.079</td>
<td>T[0.06,0.079, 0.098]</td>
</tr>
<tr>
<td>HSP</td>
<td>Heating set point (°C)</td>
<td>RES</td>
<td>26</td>
<td>U[24, 28]</td>
</tr>
<tr>
<td>CSP</td>
<td>Cooling set point (°C)</td>
<td>RES</td>
<td>26.5</td>
<td>U[25, 28]</td>
</tr>
</tbody>
</table>

Figure 2: Spatial distribution of the building stock in the Marina district
Urban building energy modeling

In this study, CityBEM was implemented as a new UBEM tool for buildings and other infrastructure (Katal et al., 2019). CityBEM is a physics-based simulation model that calculates urban thermal loads and energy use. The model takes building information, uses, and operations as inputs and is capable of running annual calculations on an hourly or shorter-time-step basis. CityBEM treats each building as a single-block air capsule with an indoor air cavity enclosed by walls and windows. The model solves the conservation of mass and energy equations to analyze building energy and indoor air temperature, improving the previous method by calculating indoor temperature using a transient heat balance equation, which is ideal for capturing transient temperature effects, such as those that may occur during power outages or emergencies. Previous studies have validated the CityBEM model (Katal et al., 2019), which provides accurate results for urban building energy analysis and is suitable for simulating various scenarios related to building energy efficiency and sustainability.

Results

CityBEM was applied as the UBEM tool to estimate the hourly and daily building cooling energy consumption in July in the Marina district of Lusail City, Qatar. Three characterization approaches were applied, the deterministic approach was considered as a baseline, and two probabilistic approaches with uniform distribution and certain distribution. Figure 1 shows the detailed hourly cooling energy profiles of the entire Marina district on July 21, considered a typical hot day. The hourly cooling energy consumption generated by the deterministic approach (method 1) was compared against the using two probabilistic approaches (methods 2 and 3) using a boxplot. For each box, the central line represents the median, while the bottom and top edges of the box indicate the 25th percentile and the 75th percentile, respectively. The whiskers indicate the maximum and minimum, not considering outliers. The boxplot shows that by considering uncertainties of inputs based on the literature and standards, the uncertainty analysis allows for obtaining a range of simulated values.

A comparison was made between the baseline method (method 1) and two probabilistic approaches (methods 2 and 3) for estimating cooling loads. The deterministic approach used values equivalent to the median value within the input range of the probabilistic approaches. The results showed that the deterministic approach consistently underestimated the cooling load. The median values of the probabilistic methods were 8% lower for the certain distribution and 13% lower for the uniform distribution. The uniform distribution produced a 5% higher cooling load compared to the certain distribution, despite having the same PDF range as the other probabilistic approach. The cooling load profile showed a large range with a uniform distribution, and the same peak hours appeared around the 12th and 13th hours of the day. The deviation from the median value varied depending on the hour of the day, with larger deviations occurring at night. The deterministic approach was close to the 75th percentile of the boxplot.

Figure 3: Hourly cooling energy profiles of the Marina district on a hot day.

Figure 4 shows the daily cooling energy of the entire Marina district for July using three different characterization approaches. The daily cooling load estimated using the deterministic approach was lower than the median value of both probabilistic approaches, with a difference of 4.3% for the certain distribution and 7.5% for the uniform distribution. The median value of the uniform distribution was 3% higher than that of the certain distribution. The range of cooling energy was greater when using the uniform distribution.

Figure 4: Daily cooling energy profiles of the Marina district in July.

Conclusions

In this study, a comparative analysis of three characterization methods was performed in UBEM archetypes development. The deterministic approach served as a baseline scenario, along with two probabilistic approaches with uniform and certain distribution. The LHS is used to propagate uncertainty in input parameters in UBEM. All the data and range are collected in the literature and standards. Then, a case study of the Marina district of Lusail City, Qatar was considered to examine the different characterization methods on district cooling energy consumption using UBEM model. Results show that the hourly and daily building cooling load using deterministic approach is smaller than the median value of both probabilistic methods. The probabilistic approach using uniform distribution bring greater uncertainties to UBEM model since the range of building cooling load is higher when using certain PDFs. The deterministic

...
approach is closer to median of probabilistic value for daily compared with hourly.

The building cooling load estimated using the deterministic approach was consistently lower than the median value obtained using both probabilistic approaches, whether on an hourly or daily basis. Although the deterministic approach utilized values that were equivalent to the median value within the input range of the probabilistic approaches. While the value range for the two probabilistic approaches was the same, the uniform distribution introduced more uncertainty to the UBEM model than the certain distribution, as it generated a wider range of cooling load values. Furthermore, the median value of the cooling load estimate was higher for the uniform distribution than the certain distribution (by 5% for hourly and 3% for daily). This study highlights the importance of considering input parameter uncertainty in UBEM to obtain reliable energy consumption estimates. Careful consideration of the distribution of input parameter is crucial in uncertainty analysis, as it can impact urban-scale building energy consumption estimates. This analysis aims to enhance our understanding of the impact of characterization methods on UBEM model and building energy consumption estimates.

Acknowledgment

This publication was made possible by NPRP grant [#NPRP11S-1208-170073] from the Qatar National Research Fund (a member of Qatar Foundation) and the Natural Sciences and Engineering Research Council of Canada under Grant [#RGPIN-2018-06734].

References


Proceedings of the 18th IBPSA Conference
Shanghai, China, Sept. 4-6, 2023
https://doi.org/10.26868/25222708.2023.1725